DNA Polymorphisms Reveal Geographic Races of Reed Canarygrass

M. D. Casler,* M. M. Phillips, and A. L. Krohn

ABSTRACT

Reed canarygrass (*Phalaris arundinacea* L.) is a cool-season perennial with a circumglobal distribution in the northern hemisphere, native to Europe, Asia, and North America. Repeated introductions of European germplasm into North America have created confusion over the origins of reed canarygrass germplasm found in wetlands, pastures, and breeding programs. The objectives of this study were to identify sources of DNA marker variation among reed canarygrass cultivars from Europe and North America and between landraces and improved cultivars from North America. Analysis of 205 reed canarygrass plants from 15 cultivars based on 102 amplified fragment length polymorphic (AFLP) DNA markers revealed two groups of cultivars. One group consisted of three closely related but geographically diverse North American landraces that were completely separated from all other plants in only two dimensions of the AFLP incidence matrix. The complete discrimination of these plants from all European plants suggests their possible origin from native North American germplasm. These results were supported by chloroplast DNA sequence analysis, which additionally revealed separation of a potential Scandinavian cytoplasmic race from the continental European cytoplasmic race. This is the strongest evidence to date suggesting that native North American reed canarygrass germplasm has been preserved within cultivars of this species.

M.D. Casler and A. Krohn, USDA-ARS, U.S. Dairy Forage Research Center, 1925 Linden Dr., Madison, WI 53706-1108; M.M Phillips, CES Program, Biological Sciences Dep., Univ. of Wisconsin-Milwaukee, 3209 N. Maryland Ave., Milwaukee, WI 53211. Received 2 Feb. 2009. *Corresponding author (michael.casler@ars.usda.gov).

Abbreviations: AFLP, amplified fragment length polymorphism; AMOVA, analysis of molecular variance; cpDNA, chloroplast DNA; PCR, polymerase chain reaction; UPGMA, unweighted pair group method with arithmetic mean.

Reed canarygrass (*Phalaris arundinacea* L.) is a cool-season perennial with a circumglobal distribution in the northern hemisphere (Anderson, 1961). It is native to Europe, Asia, and North America. Its status as a North American native was verified by herbarium samples collected in remote regions of the U.S. Pacific Northwest in the early 19th century, before European immigrants settled that region (Merigliano and Lesica, 1998). European and North American reed canarygrass herbarium samples are largely indistinguishable from each other on a phenotypic basis, indicating that European and North American strains represent different populations that cannot be reliably differentiated based on phenotype (Merigliano and Lesica, 1998).

Reed canarygrass was cultivated in Europe as early as the mid–18th century (Always, 1931). The first cultivation of reed canarygrass in North America likely occurred about the 1830s in the northeast United States and eastern Canada, near the same time that cultivation spread from Scandinavia into other parts of northern Europe (Always, 1931; Schoth, 1929). At this time, cultivation consisted of harvesting seed from native stands and planting the species in disturbed areas, largely for reclamation of peatlands and marshes.

Published in Crop Sci. 49:2139–2148 (2009).
doi: 10.2135/cropsci2009.02.0055

677 S. Segoe Rd., Madison, WI 53711 USA

Natural European strains of reed canarygrass were imported into North America sometime before 1924 and quickly dominated the marketplace (Schoth, 1929). The specific reasons for domination by European strains is unknown but may be related to seed production, the typical driver for commercialization of forage strains for which there are little or no agronomic trial data. The use of reed canarygrass in agriculture likely increased in the 1930s, as farmers and extension personnel were desperate for perennials that could withstand the severe drought that plagued the midwestern United States in the early 1930s. Reed canarygrass is one of the most drought tolerant cool-season grasses adapted to eastern North America (Wilkins and Hughes, 1932). The ability to propagate reed canarygrass by seed, sod, or stem cuttings was an additional advantage, partly responsible for the increased use of this grass in conservation programs. The results of these efforts can still be seen today in the abundance of reed canarygrass along roadsides, streambanks, and conservation strips within and between tilled fields. As such, reed canarygrass plantings used for pasture, hay production, or soil conservation may originate from European introductions, North American natives, or hybrids that result from the high levels of self-incompatibility and subsequent outcrossing within this species.

Because native and introduced strains of reed canarygrass have coexisted in North America for over 100 yr, considerable intercrossing, mixing, and migration may have occurred between native and introduced strains. The high reed canarygrass colonization rates in many wetland environments is often attributed to superior fitness of introduced strains or hybrids between introduced and native strains. While the parental genotypes of most North American reed canarygrass cultivars can generally be traced to collections made in specific pastures or hay fields, the origin of these germplasm sources remains unknown, due to the lack of phenotypic markers to distinguish North American from European germplasm. A mechanism to identify the continental origins of reed canarygrass germplasm is important for two reasons: (i) to classify germplasm used in breeding programs, and (ii) as a basis for germplasm classification and hypothesis testing of reed canarygrass populations that have colonized many North American habitats.

The objectives of this study were to identify sources of DNA marker variation among reed canarygrass cultivars from Europe and North America and between landraces and improved cultivars from North America.

MATERIALS AND METHODS

Reed Canarygrass Germplasm

Nineteen cultivars were chosen for this study, based largely on the availability of viable seed (Table 1). Improved North American cultivars were obtained directly from commercial companies as certified seed. The remaining cultivars were obtained from the USDA–ARS Germplasm Resources Information Network (http://www.ars-grin.gov/). North American cultivars were classified as landraces, collections made in the early 20th century from wetlands that had been present for many years, and improved cultivars, synthetic populations that had undergone intentional selection for improved levels of agronomic traits (Alderson and Sharp, 1994). Passport data on European cultivars are not available, but based on a long history of cultivar improvement, it is likely that all European cultivars represent some level of improvement and that they derive from European germplasm. In Europe, unlike in North America, cultivar names are typically reserved for cultivars that have received some intentional selection and breeding.

Seeds of each cultivar were germinated and established in potting mix in the greenhouse, resulting in 35 plants per cultivar. For the amplified fragment length polymorphism (AFLP) marker diversity study, a total of 15 cultivars and up to 20 plants per cultivar were included. For the chloroplast DNA (cpDNA) sequence analysis, a total of 18 cultivars, with a goal of two plants per cultivar, were included.

Table 1. Names and origins of 19 reed canarygrass cultivars included in the DNA marker diversity studies.

Cultivar name	Continent of origin	State, province, or country	Improvement status	PI no.†
AR Upland	North America	Arkansas, USA	Landrace	578790
Auburn	North America	Alabama, USA	Landrace	422031
Superior	North America	Oregon, USA	Landrace	578792
Bellevue	North America	Quebec, Canada	Improved cultivar	NA
Chiefton	North America	Iowa, USA	Improved cultivar	NA
Grove	North America	Ontario, Canada	Improved cultivar	357645
Ioreed	North America	Iowa, USA	Improved cultivar	422030
Rise	North America	Iowa, USA	Improved cultivar	578796
Rival	North America	Manitoba, Canada	Improved cultivar	NA
Vantage	North America	Iowa, USA	Improved cultivar	578794
Venture	North America	Iowa, USA	Improved cultivar	531089
Bamse	Europe	Sweden	Improved cultivar	NA
Donskoi-18	Europe	Russia	Improved cultivar	345662
Kievskij	Europe	Ukraine	Improved cultivar	505892
Lakeside LA	Europe	Hungary	Improved cultivar	587193
Motycka	Europe	Poland	Improved cultivar	272122
Nakielska	Europe	Poland	Improved cultivar	272123
Pervence	Europe	Russia	Improved cultivar	505892
Priekul'skij 15	Europe	Russia	Improved cultivar	406316

†Plant introductions available at the USDA–ARS–NPGS Germplasm Resources Information Network (http://www.ars-grin.gov/); NA = not available in the GRIN collection.

DNA Extraction

Fresh leaves (0.1–0.2 g) were ground in liquid nitrogen. Total DNA was obtained using either DNeasy Plant Extraction Kit (Qiagen, Valencia, CA) or the method of Storchova et al. (2000) and quantified by fluorometry. Resultant DNA solutions were normalized to 100 ng μL^{-1} for AFLP reactions and to 10 ng μL^{-1} for sequencing.

AFLP Reactions

DNA fragments were amplified using a modification of the procedure of Vos et al. (1995). Restriction digestion of template DNA (100 ng) and ligation to standard AFLP adapters E00 and M00 (Applied Biosystems, Foster City, CA) were performed in 10-μL reactions overnight at room temperature followed by 20 min at 80°C. The restriction–ligation mixture contained 0.05 M NaCl, 0.5 mg ml^{-1} bovine serum albumen, 2 μM *Mse*I adaptor, 0.2 μM *Eco*RI adaptor, 5 U *Mse*I (New England Biolabs, Beverly, MA), 5 U *Eco*RI (New England Biolabs), 1 U T4 DNA ligase (Invitrogen Corp., Carlsbad, CA), and 1X T4 DNA ligase buffer with ATP (Invitrogen).

Restriction–ligation products were diluted 10-fold with 0.1X TE buffer for use as template in pre-selective amplification. The reaction mixture contained 11.25μL AFLP Core Mix (Applied Biosystems), 0.5μM *Mse*I+C primer, 0.5μM *Eco*RI+A primer, and 3μL template DNA (total volume = 15 μL). Polymerase chain reaction was run on a Genemate Genius thermal cycler (Techne Ltd., Cambridge, UK) set as follows: 72°C for 2 min; 20 cycles of 94°C for 20 s, 56°C for 30 s, and 72°C for 2 min; 72°C for 2 min; and 60°C for 30 min.

Pre-selective amplification products were diluted 10-fold with 0.1X TE buffer for use as template in selective amplification. The reaction mixture contained 7.5 μL AFLP Core Mix (Applied Biosystems), 0.25 μM *Mse*I+CTT primer, 0.05 μM *Eco*RI+AGG primer 5′-labeled with WellRED D4 (Beckman-Coulter, Fullerton, CA), and 1.5 μL template DNA (total volume = 10 μL). Thermal cycling was performed as follows: 94°C for 2 min; 10 cycles of 94°C for 20 s, 66°C for 30 s, and 72°C for 2 min; 25 cycles of 94°C for 20 s, 56°C for 30 s, and 72°C for 2 min; and 60°C for 30 min.

Amplified fragment length polymorphism fragments were separated by capillary electrophoresis on a Beckman-Coulter CEQ 8000 instrument according to manufacturer's guidelines. Fragment data within the range of 50 to 450 bp were scored and analyzed by internal CEQ analysis software. A total of 132 AFLP peaks were generated from the primer set used in this study. A total of 30 peaks were excluded from this study due to low polymorphism (<1%, present in only 1 or 2 of the 205 plants). The remaining 102 peaks were present in 1.5 to 76.1% of the plants.

AFLP Statistical Analysis

Genetic distances among the 205 plants in all pairwise combinations were estimated as the complement to Jaccard's similarity coefficient (Gower, 1972). Multidimensional scaling (PROC MDS; SAS Institute, 1999) was used to organize the genetic distance matrix into two orthogonal scales to obtain a two-dimensional visual image of AFLP marker variation on an individual-plant basis. Analysis of molecular variance (AMOVA; Excoffier et al., 2005) was performed in GenAlEx6 (Peakall and Smouse, 2006) on all individuals, partitioning the genetic distance matrix into continents, cultivars within continents, and plants within cultivars. The 113 individuals from North American cultivars were analyzed in a separate AMOVA, partitioning variation into cultivar type (landrace vs. improved cultivar), cultivars within types, and plants within cultivars. Variance components were estimated by equating AMOVA mean squares to their expectations and were tested by nonparametric permutation tests (Schneider et al., 1997). Average gene diversity was computed for each cultivar using the formula of Nei (1987, p. 257). Within-cultivar mean squares were computed as described by Peakall and Smouse (2006). Mean genetic distances were computed within each cultivar and within each of the three groups of cultivars (North American landraces, North American improved cultivars, and European cultivars), using the individual genetic distance values in the 205 × 205 distance matrix. Finally, the frequency of each AFLP marker was computed for each of the 15 cultivars, from which the principal components were used in an unweighted pair group method with arithmetic mean (UPGMA) cluster analysis of the 15 cultivars. The percentage of variation associated with apparent cluster groups was determined by an *a posteriori* AMOVA.

cpDNA Sequence Analysis

Chloroplast DNA was amplified from the following 13 intergenic regions: trnL(UAA) intron, 3′trnL(UAA)-trnF(GAA) (Taberlet et al., 1991); trnH(GUG)-psbA (Hamilton 1999); trnS(GCU)-psbD, psbM-ORF29R (Saltonstall, 2001); rpL16 intron (Small et al., 1998); rpS16 intron (Shaw et al., 2005); psbJ-petA, rpl32-trnL(UAG), trnQ(UUG)-5′rps16, petL-psbE, atpI-atpH ndhA intron (Shaw et al., 2007). Reactions were performed in 6-μL volumes containing 1X JumpStart REDTaq ReadyMix (Sigma-Aldrich, St. Louis, MO), 0.2 μM each primer, 1 M betaine (Sigma-Aldrich) and 10 ng template DNA. Thermal cycling was performed on a BioRad DNA Engine Dyad thermal cycler as follows: 80°C for 5 min; 35 cycles of 95°C for 1 min, 50°C for 1 min with a ramp of 0.3°C s^{-1}, 65°C for 5 min.

Polymerase chain reaction (PCR) products were purified by adding 2 μL of 0.1 U μL^{-1} Exonuclease I (USB Corp., Cleveland, OH) and 0.1 U μL^{-1} Shrimp Alkaline Phosphatase (USB Corp.) and incubating 30 min at 37°C followed by 20 min at 80°C. Cycle sequencing was performed using a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) according to the method of Platt et al. (2007) and resolved on an ABI 3730 Genetic Analyzer (Applied Biosystems). Sequences were aligned in BioEdit 7 (Hall, 1999) using the CLUSTAL W algorithm (Thompson et al., 1994). Terminal coding regions of the PCR amplicons were excluded from the contigs following comparison to the annotated chloroplast genomes of *Hordeum vulgare*, NC 008590, and *Agrostis stolonifera*, NC 008591 (Saski et al., 2007). Unique sequences generated during this study were deposited in GenBank with accession numbers FJ590473–FJ590497.

Seven of the 13 regions were informative with at least one polymorphic locus. Those seven regions accounted for 4789 bp of annotated DNA sequence and contained a total of 27 polymorphic sites: 16 substitutions and 11 indels. The seven informative regions were sequenced in each individual to an average depth of 2.7. Polymorphic sites were sequenced between 1 and 11 times to an average depth of 2.2. Data for each site were converted into an arbitrary numerical variable with values of 0 or 1 and all 27 positions were used in an UPGMA cluster analysis of the 35 plants.

RESULTS

Analysis of molecular variance revealed two levels of structure among the reed canarygrass cultivars. First, AFLP marker

Table 2. Analyses of molecular variance (AMOVA) for 205 reed canarygrass plants originating in either North America (NA) or Europe and for 113 plants originating from two improvement-status levels for North American cultivars (landraces vs. improved cultivars), based on 102 amplified fragment length polymorphic (AFLP) DNA markers.

Source of variation	df	Variance component	Percentage of variation	*P* value
AMOVA of all 205 plants				
Continent	1	0.43	5.0	<0.0001
Cultivar/continent	13	1.07	12.5	<0.0001
Plants/cultivar/continent	190	7.11	82.5	<0.0001
AMOVA of 113 NA plants				
Improvement status (IS)	1	1.65	19.6	<0.0001
Cultivar/IS	6	0.24	2.9	0.0081
Plants/cultivar/IS	105	6.54	77.5	<0.0001

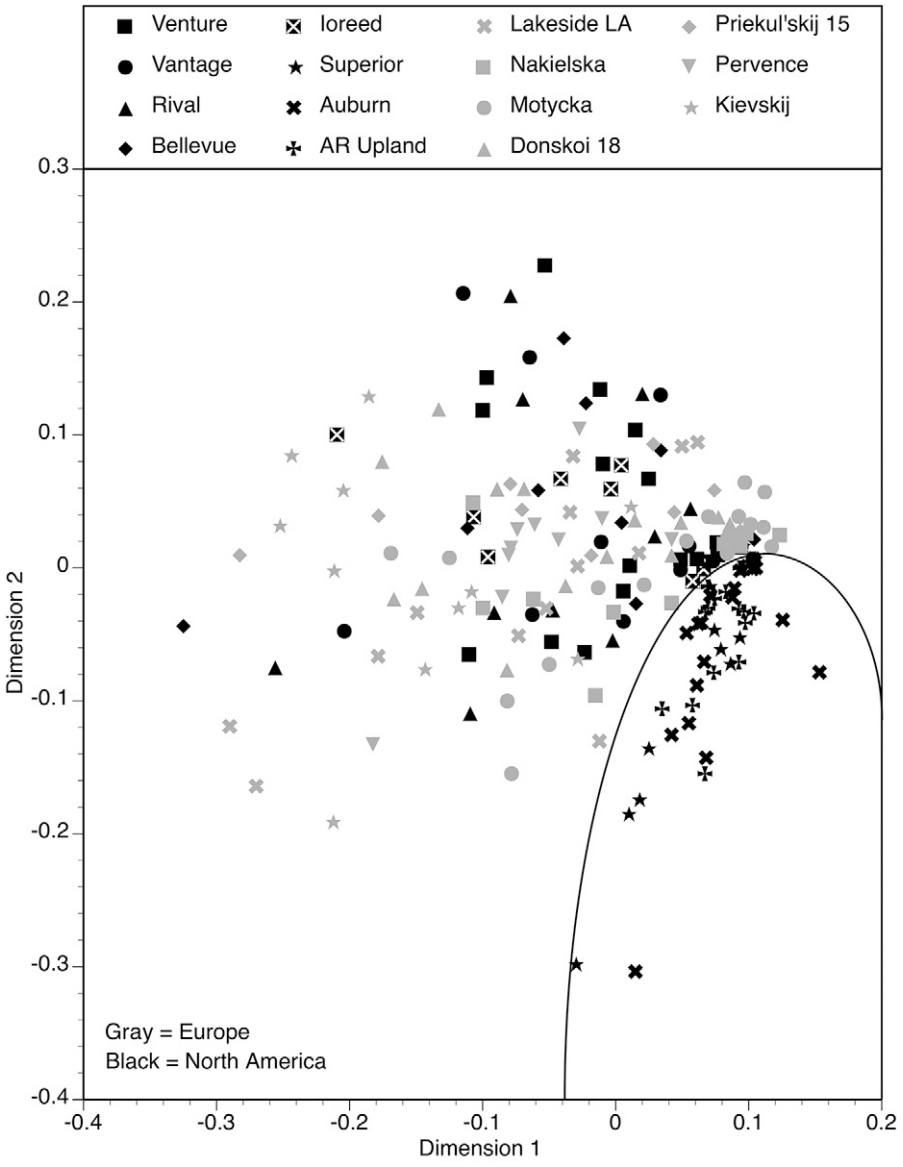

Figure 1. Two-dimensional scale plot of 205 reed canarygrass plants evaluated for 102 amplified fragment length polymorphic (AFLP) DNA markers. The ellipse divides all plants of the three North American landraces (AR Upland, Auburn, and Superior) from all other plants.

variation was significant between the North American and European continents, accounting for 5% (Table 2) of the total marker variation and 28% (5.0/17.5) of the marker variation among cultivars. Mean genetic distance of plants originating on two different continents was 0.172 ± 0.004, while the mean genetic distance of plants originating on the same continent was 0.154 ± 0.003 (data not shown). These two distance coefficients were also significantly different by *t* test ($P < 0.0001$). North American and European cultivars were not distinctly separated from each other in two dimensions (Fig. 1) or on the basis of a cluster dendrogram that utilized the full dimensionality of the AFLP marker incidence matrix (Fig. 2). The cluster dendrogram revealed four groups that when analyzed by AMOVA, accounted for 71.0% of AFLP marker variability among the 15 cultivars. The largest of these groups contained six of seven European cultivars and four of five improved North American cultivars. The European cultivar Kievskij and the North American cultivar Rival each formed monotypic clusters at the level of 71% of the cultivar variance explained (four clusters). Neither of these two cultivars could be differentiated from the other modern cultivars on the basis of two dimensions (Fig. 1), indicating relatively complex differentiation on the basis of numerous AFLP markers.

Second, the comparison of landraces versus improved cultivars accounted for 19.6% of the total marker variation among plants from North American cultivars (Table 2) and 87% (19.6/22.5) of the marker variation among the eight North American cultivars. The three North American landraces—AR Upland, Auburn, and Superior—were completely distinct from all European cultivars and all improved North American cultivars (Fig. 1). Only two dimensions of the AFLP marker incidence matrix were required to demonstrate this level of differentiation. These three cultivars also formed the most closely related group within the cluster dendrogram, which utilized all information contained within the AFLP marker incidence matrix (Fig. 2).

Marker variances (mean squares) within cultivars were larger for improved cultivars compared with landraces originating in North America (Table 3). Pooled across cultivars with the three groups, mean squares were 7.22 for European cultivars, 7.34 for improved North American cultivars, and 4.64 for landraces. Mean genetic distances of plants within cultivars tended

to be larger for improved cultivars compared to landraces. Taken across all plants of the three groups, mean genetic distance of plants within groups were: 0.167 ± 0.001 for European cultivars, 0.160 ± 0.002 for improved North American cultivars, and 0.097 ± 0.002 for North American landraces. Mean distances of North American landraces were significantly different from each of the other two mean values by *t* test ($P < 0.0001$). Similarly, mean marker diversity was 0.330 ± 0.008 for European cultivars, 0.337 ± 0.009 for improved North American cultivars, and 0.270 ± 0.015 for North American landraces. Mean marker diversity of North American landraces was significantly less than for each of the other two groups by *t* test ($P < 0.0001$).

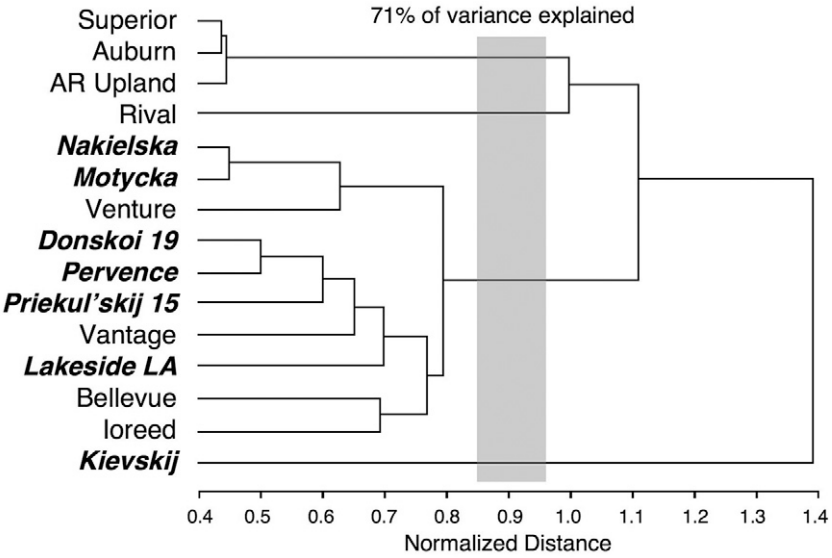

Figure 2. Cluster dendrogram of 15 reed canarygrass cultivars, based on unweighted pair group method with arithmetic mean clustering of 102 amplified fragment length polymorphic (AFLP) DNA marker frequencies within each cultivar. Bold and italic names are European cultivars. Vertical gray bar indicates identity of four clusters that explain 71% of the AFLP marker variation among the 15 cultivars.

Most of the AFLP marker variation was found among plants within cultivars, reflecting the highly self-incompatible reproductive system of reed canarygrass that promotes a high level of cross-pollination and perhaps multiple origins of parental genotypes. This result aligns well with results from other cross-pollinated perennial grasses for both dominant and codominant marker systems (Xu et al., 1994; Knapp and Rice, 1996; Fjellheim and Rognli, 2005). Similarly, a study of AFLP marker variability on *P. aquatica* L., the most closely related species to reed canarygrass, found 74% of marker variation to be within populations (Mian et al., 2005).

Based on 27 single-nucleotide sites within the seven polymorphic cpDNA regions, a total of 15 haplotypes were identified (Table 4). The 15 haplotypes fell into three groups as indicated by sequence observations and cluster analysis (Fig. 3). The continental European group was the most diverse, containing plants of six European cultivars and three North American cultivars and a total of nine haplotypes. Considerable haplotype homology was observed within cultivars, but clearly some cultivars, such as 'Donskoi-18', 'Kievskij', 'Bellevue', and 'Chiefton', originated from multiple germplasm sources, as indicated by two distinct haplotypes. Rival represents a more diverse parentage, derived from 'Grove', which originated from the putative Scandinavian cytoplasmic race, and Ottawa Synthetic C, which must have originated from the continental European cytoplasmic race. The continental European group included polymorphisms at 10 of the 27 single-nucleotide sites.

One European cultivar, 'Bamse' from Sweden, had a unique haplotype, shared only by one plant from the North American cultivar Venture (Table 4). The remaining plants within this group, six plants from three North American cultivars and one plant from the Polish cultivar Motycka, also had a unique haplotype that differed from the Bamse haplotype at only three single-nucleotide sites. These results suggest the presence of a possible Scandinavian cytoplasmic race of reed canarygrass that has been historically important to North American reed canarygrass breeding programs (Fig. 3). Plants of the putative Scandinavian cytoplasmic race were highly uniform, containing polymorphisms at only three of the 27 cpDNA positions.

The third group contained all plants of the three North American landraces, including two haplotypes, and both

Table 3. Measures of amplified fragment length polymorphic (AFLP) DNA marker diversity and numbers of plants sampled (*n*) within 15 reed canarygrass cultivars.

Cultivar	Origin†	*n*	Marker diversity‡	Mean square§	Mean distance¶
AR Upland	NA-LR	19	0.245 ± 0.028	3.67	0.071 ± 0.003
Auburn	NA-LR	20	0.250 ± 0.025	5.13	0.100 ± 0.005
Superior	NA-LR	14	0.309 ± 0.023	6.33	0.123 ± 0.009
Bellevue	NA-IC	9	0.341 ± 0.018	9.39	0.182 ± 0.014
loreed	NA-IC	9	0.372 ± 0.024	7.25	0.141 ± 0.010
Rival	NA-IC	12	0.342 ± 0.022	9.05	0.176 ± 0.008
Vantage	NA-IC	10	0.356 ± 0.020	8.56	0.166 ± 0.010
Venture	NA-IC	20	0.281 ± 0.022	6.89	0.134 ± 0.005
Donskoi-18	EU	15	0.337 ± 0.022	7.58	0.147 ± 0.005
Kievskij	EU	11	0.320 ± 0.018	10.07	0.196 ± 0.006
Lakeside LA	EU	15	0.312 ± 0.021	9.51	0.185 ± 0.008
Motycka	EU	17	0.319 ± 0.023	6.85	0.133 ± 0.006
Nakielska	EU	13	0.340 ± 0.027	5.62	0.109 ± 0.007
Pervence	EU	12	0.321 ± 0.023	7.05	0.137 ± 0.007
Priekul'skij 15	EU	9	0.370 ± 0.019	8.69	0.169 ± 0.012

†NA = North America, EU = Europe, LR = landrace, IC = improved cultivar.

‡Mean marker diversity as defined by Nei (1987).

§Within-cultivar mean square (Peakall and Smouse, 2006).

¶Mean genetic distance computed from all pairs of plants within each cultivar.

Table 4. Single nucleotide polymorphism chloroplast DNA (cpDNA) markers from sequence analysis of 35 reed canarygrass plants representing 18 cultivars.

Group cultivar (plant no.)	Origin	cpDNA region, number of base pairs, and base-pair position number†																
		3′trnL(UAA)-trnF(GAA) (338 bp)	trnH(GUG)-psbA (254 bp)			atpI-atpH (567 bp)	psbJ-petA (695 bp)			rpL16 intron (976 bp)		rpS16 intron (827 bp)		petL-psbE (1132 bp)		
		110–112	133	156	188–189	29–30	323–325	549–554	649	142	659	165	463	653	656	817–818
Scandinavian group																
Bamse (1)	Scandinavia	GG–	C	A	TC	– –	– – –	TTTTTT	A	A	C	A	T	G	G	– –
Bamse (2)	Scandinavia	GG–	C	A	TC	– –	– – –	TTTTTT	A	A	C	A	T	G	G	– –
Venture (3)	North America	GG–	C	A	TC	– –	– – –	TTTTTT	A	A	C	A	T	G	G	– –
Vantage (1)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Vantage (2)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Rise (2)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Rise (3)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Grove (1)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Grove (2)	North America	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Motycka (3)	Continental Europe	GG–	C	A	GA	– –	T – –	TTTTTT	A	A	C	A	T	G	G	– –
Continental European group																
Donskoi-18 (2)	Continental Europe	– – –	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	G	G	T–
Donskoi-18 (5)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
Kievskij (3)	Continental Europe	GG–	C	A	TC	– –	TT–	AAAAAA	A	A	C	A	T	G	–	T–
Kievskij (4)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
Lakeside LA (1)	Continental Europe	GGG	C	A	GA	– –	TT–	AAAAAA	A	A	T	A	T	G	G	T–
Lakeside LA (2)	Continental Europe	GGG	C	A	GA	– –	TT–	AAAAAA	A	A	T	A	T	G	G	T–
Motycka (2)	Continental Europe	– – –	C	A	TC	– –	TT–	AAAAAA	A	A	C	A	T	G	G	T–
Nakielska (2)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	C	G	G	T–
Nakielska (3)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	C	G	G	T–
Priekul'skij-15 (1)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
Priekul'skij-15 (2)	Continental Europe	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
Rival (1)	North America	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	G	G	– –
Bellevue (1)	North America	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	G	G	T–
Bellevue (2)	North America	– – –	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	G	G	T–
Chiefton (2)	North America	– – –	C	A	TC	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
Chiefton (5)	North America	GG–	C	A	GA	– –	TT–	AAAAAA	A	A	C	A	T	A	G	T–
North American group																
AR Upland (1)	North America	G – –	G	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
AR Upland (2)	North America	G – –	G	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
AR Upland (3)	North America	G – –	G	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
Superior (10)	North America	G – –	C	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
Superior (12)	North America	G – –	C	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
Auburn (1)	North America	G – –	C	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
Auburn (2)	North America	G – –	C	A	GA	TT	TTT	AAAAAA	C	G	C	A	T	G	G	TT
Ioreed (1)	North America	G – –	C	A	TC	TT	TTT	AAAAAA	C	A	C	A	T	G	G	T–
Ioreed (2)	North America	G – –	C	T	TC	TT	TT–	AAAAAA	A	A	C	C	T	G	G	T–

†cpDNA sequences: 5′trnL(UUA)-3′trnL(UAA), 3′trnL(UAA)-trnF(GAA) (Taberlet et al., 1991); trnH(GUG)-psbA (Hamilton, 1999); trnS(GCU)-psbD, psbM-ORF29R (Saltonstall, 2001); rpL16 intron (Small et al., 1998); rpS16 intron (Shaw et al., 2005); psbJ-petA, rpl32-trnL(UAG), trnQ(UUG)-5′rps16, petL-psbE, atpI-atpH, ndhA intron (Shaw et al., 2007).

plants of the old cultivar Ioreed, including two additional haplotypes. All plants of the three North American landraces were completely distinguishable from all other plants by a minimum of six single-nucleotide polymorphisms contained in four of the seven cpDNA regions, atpI-atpH, psfJ-petA, rpL16 intron, and petL-psbE (Table 4). Plants of AR Upland were separated from plants of Auburn and Superior by only one polymorphic site. The two plants of Ioreed appear to be intermediate in cpDNA sequence polymorphism to the North American and continental European groups but are most closely associated with the North American group. Ioreed originated from multiple sources, including germplasm from Germany (14%), Oregon (7%), and the central United States (79%) (Alderson and Sharp, 1994). Plants within the putative North American cytoplasmic race contained polymorphisms at 9 of 27 single-nucleotide sites, but most of these polymorphisms were associated with Ioreed versus the landraces or diversity within Ioreed.

DISCUSSION

The knowledge that reed canarygrass is native to North America and that early agricultural use of this species was largely based on European accessions in the early 20th century has created considerable uncertainty over the origins of modern reed canarygrass populations. This uncertainty extends across a wide range of habitats and uses, including old pastures and hayfields of unknown origin, modern cultivars that largely derive from these old pastures, and most reed canarygrass populations that have colonized wetlands across North America. The combination of ancestral ambiguity with the clear demonstration of an invasive hybrid of *Phragmites* (Saltonstall, 2002) has led to suggestions that European populations per se or European × native hybrids are more aggressive than native populations and, as such, are responsible for the increased aggressiveness of reed canarygrass as a wetland invader during the latter half of the 20th century (Merigliano and Lesica, 1998). Numerous references to this hypothesis have elevated it to the level of commonly accepted dogma (e.g., Lindig-Cisneros and Zedler, 2002; Maurer and Zedler, 2002; Spuher, 1994), despite lack of scientific support.

The results of AFLP marker analyses suggest that cultivars can be broken into two clear groups based on nuclear DNA: a small group of North American landraces and all others. The three North American landraces available to us at the initiation of this study, those deriving from Arkansas, Alabama, and Oregon, are quite possibly derived from native North American germplasm. Of course, these landraces may originate from European germplasm that was quite common in North America during the late 19th and early 20th centuries (Schoth, 1929). However, the complete separation of North American landraces originating from two highly disparate regions of the United States—Superior from Oregon, Auburn and AR Upland from the southeast—from all European cultivars, which, themselves, originated from a wide region of eastern and northern Europe suggests differential origins of these three landraces from all other germplasm in this study. Reed canarygrass is far less common in Arkansas or Alabama than it is in more northerly regions of the United States, whether in wetlands, conservation plantings, or pastures. Agricultural activities involving reed canarygrass would be far less common in this region than in the northern United States. Thus, old landraces from this region, which represents the southern border of reed canarygrass adaptation in the United States (Sheaffer and Marten, 1995), would be the most logical source of native reed canarygrass germplasm among cultivars.

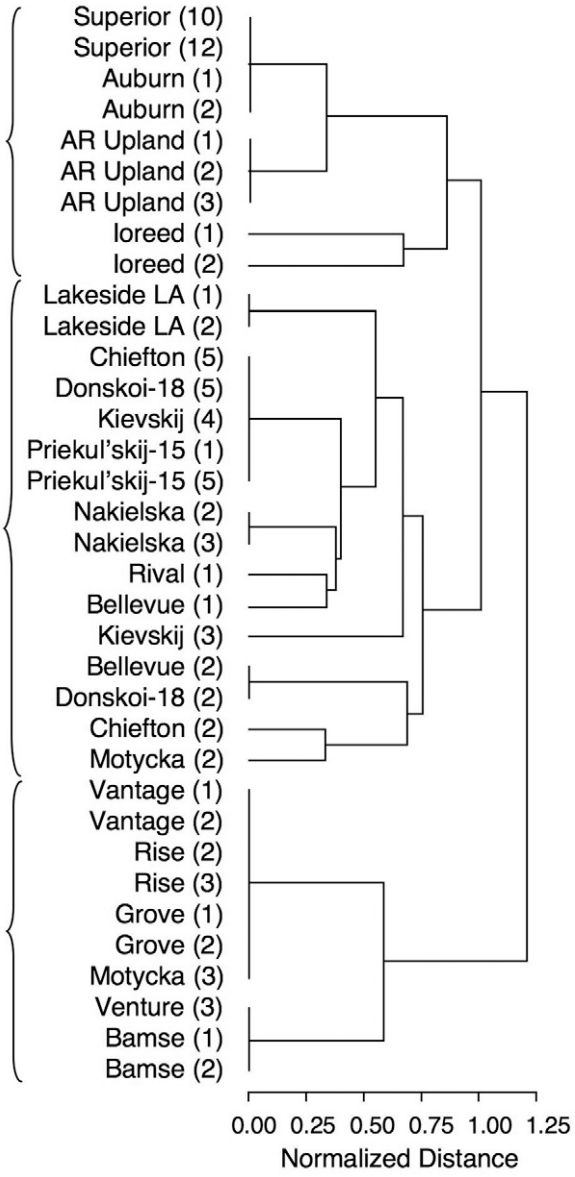

Figure 3. Cluster dendrogram of 35 reed canarygrass plants, based on unweighted pair group method with arithmetic mean clustering of 27 single-nucleotide polymorphism cpDNA markers. Numbers in parentheses are arbitrary numbers assigned to different plants within each cultivar.

Improved North American cultivars could not be separated from the European cultivars based on nuclear DNA, with the exception of Rival, which appeared to be a possible intermediate form. Most reed canarygrass cultivars bred in North America derive from plants collected in old pastures (Alderson and Sharp, 1994). Since the 1970s, these plants have all been screened for alkaloid type and concentration to create cultivars acceptable for use in managed rotational grazing systems (Coulman et al., 1976; Marten, 1989). If some of these pastures originated from native reed canarygrass stands that were managed as seed orchards in the late 19th or early 20th centuries, it is quite possible for cultivars such as Rival to be derived from both native and European germplasm. The use of native- and European-derived plants as parents of one synthetic cultivar would result in a mixture of native, European, and hybrid plants, resulting in a combination of native, European, and mixed AFLP haplotypes after three to four generations of seed multiplication. Indeed, the parents of Rival derive from two distinct sources: the cultivar Grove and the germplasm Ottawa Synthetic C. Grove derives from purportedly "native" collections in eastern Canada (Alderson and Sharp, 1994), which might include true native germplasm. However, the cpDNA analyses suggest that Grove derives from the putative Scandinavian cytoplasmic race, probably as selections from old fields originally established from seed introduced by Scandinavian immigrants. Mixed pedigrees are common in breeding forage cultivars, as evidenced by the high level of cpDNA polymorphism within the Polish cultivar Motycka, as breeders will frequently use multiple germplasm sources in the parentage of a single cultivar.

The lack of separation among improved North American cultivars and the European cultivars creates the very tempting urge to conclude that the invasion of reed canarygrass into North American wetlands during the latter half of the 20th century was a direct result of the introduction and heavy use of European germplasm in agriculture. This is the reason that Lavergne and Molofsky (2004) called for a ban on reed canarygrass breeding. It must be strongly emphasized that no evidence exists to support this hypothesis, which derives largely by extension and extrapolation of the verified introduced "invasive" genotype of *Phragmites australis* (Saltonstall, 2002). Numerous field trials of North American and European cultivars have failed to demonstrate any difference in vegetative fitness or potential invasiveness between North American and European cultivars (Brummer and Casler, unpublished data, 1985–2001). Likewise, reed canarygrass populations from pastures and colonized wetlands did not differ in phenotypic traits that relate to plant vigor (Gifford et al., 2002). Studies of reed canarygrass germplasm from around the world have verified that seed shattering, rhizomatous spreading, and long-term persistence in wet or waterlogged soils are universal traits of this species. Rather, "guilt by association" is more likely, as there is overwhelming evidence that reed canarygrass colonization of wetlands occurs by the simultaneous movement of seeds, sediments, and nutrients into wetlands, overwhelming or burying native vegetation, and providing an extremely favorable environment for seedling recruitment of reed canarygrass (Green and Galatowitsch, 2001, 2002; Maurer and Zedler, 2002; Maurer et al., 2003; Perry et al., 2004). European-derived populations may be much more frequent in wetlands than North American populations, but it is highly unlikely that this is due to genetic differences in fitness or invasiveness or to hybrid vigor between European and North American populations. Management of old reed canarygrass stands is often very lax, sometimes due to waterlogging of lowland soils, allowing seed ripening, shattering, and transport on surface waters. Populations that dominate wetlands are likely similar in genetic makeup and origin to those that dominate old agricultural fields that are frequently allowed to produce ripe seed.

By all measures used in this study, the three North American landraces that represent putative native germplasm had significantly less genetic variability, both within and among cultivars, compared with all improved cultivars from either North America or Europe. This suggests the existence of a founder effect, resulting from the migration of reed canarygrass from Europe or Asia into North America during one or more recent interglacial periods. While this species is considered to be native to post-Columbian North America, its tenure as a North American native appears to be considerably shorter than its European life history. North American reed canarygrass founder populations likely accumulated thousands of mutations that eventually led to haplotypes, both nuclear and chloroplast, distinct from their European ancestors, as observed in the three North American landraces. Most of these mutations had little effect on fitness, morphological, or life history traits, leaving the species essentially unaltered, on a phenotypic basis, from its migratory forms, but with reduced genetic variability resulting from the genetic bottleneck. Our results suggest that we have detected a sufficiently large number of those mutations to discriminate descendants of these North American founders from the descendants of the original European populations.

The more recent introduction of European reed canarygrass germplasm to North America likely began in the late 19th century and continued through the early 20th century, involving many separate immigration events from throughout its European range (Always, 1931; Schoth, 1929). Multiple immigration events, combined with migration across North America, including both human-induced seed dispersal and wind-aided pollen flow, have resulted in significant intercrossing and mixing of North American and European gene pools. This can be observed in cultivars such as Ioreed, containing the North American cytoplasmic haplotype and European nuclear haplotype, and Rival, possessing mixed Scandinavian and continental European

ancestry. Polymorphic patterns observed within Ioreed clearly suggest a mixture of North American and European lineages. These results support the hypothesis that recent immigrations of European reed canarygrass populations have overcome the genetic bottleneck created during the initial colonization of North America by this species (Lavergne and Molofsky, 2007). Increased genetic variation during the 20th century has likely led to increased phenotypic plasticity and greater potential for natural selection in a diverse range of habitats and environments.

CONCLUSIONS

The results of this study are the first evidence suggesting that native reed canarygrass germplasm has been preserved within cultivars of this species. However, before these results are taken for granted or accepted as fact, they should be validated on additional DNA samples, including additional North American landraces, additional improved cultivars from North America, and European accessions that have not undergone selection and breeding. Should this conclusion be validated, these native North American landraces will become a valuable germplasm resource for hypothesis-driven experiments to determine the mechanisms behind the late-20th-century invasion of reed canarygrass into many North American wetlands, whether genetics has played a role in this phenomenon either as a natural consequence of mutation, migration, selection, and drift, or as a result of human-driven selection, breeding, and dissemination of germplasm. While it is becoming increasingly clear that recent introductions of European germplasm have increased genetic variability within North American reed canarygrass germplasm samples, the mechanism remains cloudy. Identification of native strains will allow for hypothesis tests related to fitness differentials among native North American strains, European strains, and North American × European hybrids and will allow a description and quantification of the presence of North American versus European lineages within habitats that are already highly colonized by reed canarygrass.

Acknowledgments

Mention of a trademark or brand name does not imply endorsement of a product over any other product by USDA-ARS or the authors of this paper.

References

Alderson, J., and W.C. Sharp. 1994. Grass varieties in the United States. USDA–SCS Agric. Handb. 170. U.S. Gov. Print. Office, Washington, DC.

Always, F.J. 1931. Early trials and use of reed canary grass as a forage plant. J. Am. Soc. Agron. 23:64–66.

Anderson, D.E. 1961. Taxonomy and distribution of the genus *Phalaris*. Iowa State J. Sci. 36:1–96.

Coulman, B.E., D.L. Woods, and K.W. Clark. 1976. Identification of low alkaloid genotypes of reed canarygrass cultivars. Can. J. Plant Sci. 56:837–845.

Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinform. Online 1:47–50.

Fjellheim, S., and O. Rognli. 2005. Genetic diversity within and among Nordic meadow fescue (*Festuca pratensis* Huds.) cultivars determined on the basis of AFLP markers. Crop Sci. 45:2081–2086.

Gifford, A.L.S., J.B. Ferdy, and J. Molofsky. 2002. Genetic composition and morphological variation among populations of the invasive grass, *Phalaris arundinacea*. Can. J. Bot. 80:779–785.

Gower, J.C. 1972. Measures of taxonomic distance and their analysis. p. 1–24. *In* J.S. Weiner and J. Huizinga (ed.) The assessment of population affinities in man. Clarendon Press, Oxford, UK.

Green, E.K., and S.M. Galatowitsch. 2001. Differences in wetland plant community establishment with additions of nitrate-N and invasive species (*Phalaris arundinacea* and *Typha* × *glauca*). Can. J. Bot. 79:170–178.

Green, E.K., and S.M. Galatowitsch. 2002. Effects of *Phalaris arundinacea* and nitrate-N addition on the establishment of wetland plant communities. J. Appl. Ecol. 39:134–144.

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41:95–98.

Hamilton, M.B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8:521–523.

Knapp, E.E., and K.J. Rice. 1996. Genetic structure and gene flow in *Elymus glaucus* (blue wildrye): Implications for native grassland restoration. Restor. Ecol. 4:1–10.

Lavergne, S., and J. Molofsky. 2004. Reed canary grass (*Phalaris arundinacea*) as a biological model in the study of plant invasions. Crit. Rev. Plant Sci. 23:415–429.

Lavergne, S., and J. Molofsky. 2007. Increased genetic variation and evolutionary potential drive the success of an invasive grass. Proc. Natl. Acad. Sci. USA 104:3883–3888.

Lindig-Cisneros, R., and J.B. Zedler. 2002. Relationships between canopy complexity and germination microsites for *Phalaris arundinacea* L. Oecologia 133:159–167.

Marten, G.C. 1989. Breeding forage grasses to maximize animal performance. p. 71–104. *In* D.A. Sleper et al. (ed.) Contributions from breeding forage and turf grasses. CSSA Spec. Publ. 15. CSSA, Madison, WI.

Maurer, D.A., R. Lindig-Cisneros, K.J. Werner, S. Kercher, R. Miller, and J.B. Zedler. 2003. The replacement of wetland vegetation by reed canarygrass (*Phalaris arundinacea*). Ecol. Restoration 21:116–119.

Maurer, D.A., and J.B. Zedler. 2002. Differential invasion of a wetland grass explained by tests of nutrients and light availability on establishment and clonal growth. Oecologia 131:279–288.

Merigliano, M.E., and P. Lesica. 1998. The native status of reed canarygrass (*Phalaris arundinacea* L.) in the inland northwest, USA. Nat. Areas J. 18:223–230.

Mian, M.A.R., J.C. Zwonitzer, Y. Chen, M.C. Saha, and A.A. Hopkins. 2005. AFLP diversity within and among harding-grass populations. Crop Sci. 45:2591–2597.

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.

Peakall, R., and P.E. Smouse. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol. Ecol. Notes 6:288–295.

Perry, L.G., S.M. Galatowitsch, and C.J. Rosen. 2004. Competitive control of invasive vegetation: A native wetland sedge

suppresses *Phalaris arundinacea* in carbon-enriched soil. J. Appl. Ecol. 41:151–162.

Platt, A.R., R.W. Woodhall, and A.L. George, Jr. 2007. Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol. Biotechniques 43:58–62.

Saltonstall, K. 2001. A set of primers for amplification of non-coding regions of chloroplast DNA in the grasses. Mol. Ecol. Notes 1:76–78.

Saltonstall, K. 2002. Cryptic invasion of a non-native genotype of the common reed, *Phragmites australis*, into North America. Proc. Natl. Acad. Sci. USA 99:2445–2449.

SAS Institute. 1999. SAS/STAT user's guide, Version 7-1, 3415. SAS Inst., Cary, NC.

Saski, C., S.B. Lee, S. Fjellheim, C. Guda, R.K. Jansen, H. Luo, J. Tomkins, O.A. Rognli, H. Daniell, and J.L. Clarke. 2007. Complete chloroplast genome sequences of *Hordeum vulgare, Sorghum bicolor*, and *Agrostis stolonifera*, and comparative analyses with other grass genomes. Theor. Appl. Genet. 115:571–590.

Schneider, S., J.M. Kueffer, D. Roessli, and L. Excoffier. 1997. Arlequin ver. 1.1: A software for population genetic data analysis. Available at http://lgb.unige.ch/arlequin/ (verified 21 Aug. 2009). Genetics and Biometry Laboratory, Univ. of Geneva, Geneva, Switzerland.

Schoth, H.A. 1929. Reed canary grass. Farmers Bull. 1602. U.S. Gov. Print. Office, Washington, DC.

Shaw, J., E.B. Lickey, J.T. Beck, S.B. Farmer, W. Liu, J. Miller, K.C. Siripun, C.T. Winder, E.E. Schilling, and R.L. Small. 2005. The tortoise and the hare: II. Relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92:142–166.

Shaw, J., E.B. Lickey, E.E. Schilling, and R. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare: III. Am. J. Bot. 94:275–288.

Sheaffer, C.C., and G.C. Marten. 1995. Reed canarygrass. p. 335–343. *In* R.F Barnes, D.A. Miller and C.J. Nelson (ed.) Forages. Vol. 1. An introduction to grassland agriculture. Iowa State Univ. Press, Ames.

Small, R.L., J.A. Ryburn, R.C. Cronn, T. Seelman, and J.F. Wendel. 1998. The tortoise and the hare: Choosing between non-coding plastome and nuclear *Adh* sequences for phylogenetic reconstruction in a recently diverged plant group. Am. J. Bot. 85:1301–1315.

Spuher, D.R. 1994. Low plant diversity found in communities dominated by reed canary grass (*Phalaris arundinacea*). p. 43–47. *In* D.C. Hartnett (ed.) Proc. of the 14th North American Prairie Conf.: Prairie Biodiversity. 12–16 July 1994. Kansas State Univ., Manhattan.

Storchova, H., R. Hrdlickova, J. Chrtek, Jr., M. Tetera, D. Fitze, and J. Fehrer. 2000. An improved method of DNA isolation from plants collected in the field and conserved in NaCl/CTAB solution. Taxonomy 49:79–84.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105–1109.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. 23:4407–4414.

Wilkins, P.S., and H.D. Hughes. 1932. Agronomic trials with reed canary grass. J. Am. Soc. Agron. 24:18–28.

Xu, W.W., D.A. Sleper, and G.F. Krause. 1994. Genetic diversity of tall fescue germplasm based on RFLPs. Crop Sci. 34:246–252.